

My Project Profile

I have worked in a very broad scope of fields like Data Science, Computer Vision, Deep Learning, Sequence Models, Neural Networks, Data Clustering, Cryptography, Making a Debugger, and Web Development

April-2020 and Ongoing

Build two models from Convolutional Neural Networks for the classification of Images from MNIST dataset using the TensorFlow in my Coursework of TensorFlow Specialisation

Completed a Course Work of Sequential Analysis from Andrew-NG through Coursera and developed several mini projects like Trigger Word Detection, Removing Word Biasing

Platform used for development is Google Colab and Jupyter Notebook

Working language : Python

February-2020 and Ongoing

I am working with the Malaysian Company Skyhive as a Software Developer. The project we are working is a way to shift almost all the offline work of medical fields worker or clinics to an online platform.

In the initial phase of the internship, i worked as a front-end developer but now i am one of the lead of the software.

The platform we are using for communication is Discord and the development environment is GitHub and Visual Studio Code

Working language : JavaScript, HTML-CSS, JQuery, PHP

January-2020 and Ongoing

Done several mini projects in the field of Data Science and Computer Vision along with the specialisation Course of John Hopkins University from Coursera (it's ongoing). My Mini projects includes,

- Detecting the word on a piece of multi column newspaper and printing out all the faces (Computer Vision)
- Reading and Manipulating Data from various places like from SQL, API's and HTML itself (Data Science)
- Learned and converted many type of Data to a simpler form known as Tidy Data (Data Science)

Platform used for development is RStudio, Google Colab and Jupyter Notebook

Working language : R and Python

August-2019 to October-2019

Developed a complete Website (including the front-end and back-end) for Hostel Administration. Some highlights of the Website

- There are three actors of the website (Warden, Council Member, Students)
- Automatic generation of Mail's for Complaints and Leaves
- Website is secured from cross site scripting and SQL injection attack

Website link : iiitnhostel.tk

Working language : JavaScript ,HTML-CSS, JQuery ,Bootstrap, PHP

September-2019

Developed and trained the Single layer and Multi Layer Perceptron Neural Networks with an accuracy of 80% for Single layer and 50% for Multi-layer, as a part of the academic project

Working language : Python

May-2019 to July-2019

This is my internship period with International Institute of Information Technology, Hyderabad and during this timeline I developed many projects like MIPS Assembly language Debugger, Simulation of Triple DES Cryptography Algorithm and the Simulation of two Data Clustering Algorithms (K means and MST Clustering).

During the initial phase of the Internship, I have worked as a Software Developer but became a mentor in the last phase. The Platform we used for communication is Slack and for development we used GitHub and Visual Studio Code

Working language : JavaScript ,HTML-CSS, JQuery ,Bootstrap, PHP

The project that I liked the most

The project that I liked the most is the Data Science mini projects that I worked upon during my coursework.

Project: Conversion to Tidy Data

The goal of the project is to convert the Raw Data or Data in other form to a suitable format which further can be used for processing the queries

There are lot's and lot's of data being generated and are available as many applications generate data like various API's or Website's Database, but the problem is the format of the data available is not the one that can be processed easily like JSON data or XML data formats are very common but to apply Data Analytics on these format is not an easy task. And hence it's required to pre-process the data and made it available in a Tidy data format.

The most brain-storming part of the problem was to decide which columns to keep and what to merge, so I just followed the guidelines for the Tidy data set to make that decision

I don't have many files locally, as i worked online and submitted. But i attached some of them

```
CountryCode,Long Name,Income Group,Region,Lending Category,Other groups,Currency Unit,
Latest population census,Latest household survey,Special notes,National accounts
base year,National accounts reference year,system of National Accounts,SNA price
valuation,alternative conversion factor,PPP survey year,balance of Payments manual
use,External debt Reporting status,system of trade,Government Accounting concept,IMR
data dissemination standard,Source of most recent income and expenditure data,vital
registration complete,Latest agricultural census,Latest industrial data,Latest trade
data,Latest water withdrawal data,2-alpha code,wb-2 code,table name,short name

AWA,Aruba,High Income: nonOECD,Latin America & Caribbean,,Aruban florin,2000,,1995
,,,,,,Special,,,,,,2008,,AW,AW,Aruba,Aruba

ADO,Principality of Andorra,High Income: nonOECD,Europe & Central Asia,,Euro
Register based,,,,,,General,,,Yes,,,2006,,AO,AO,Andorra,Andorra

AFG,Islamic State of Afghanistan,Low Income,South Asia,IDA,WFP,afghan-afghani,1979
,"MICS, 2008",Fiscal year end: March 20; reporting period for national accounts data
FY..2002/2001,,,VAB,,,,Actual,General,consolidated,GDDS,,,,2008,2000,AF,AF,Afghanis
an,Afghanistan

AGO,People's Republic of Angola,Lower middle Income,Sub-Saharan Africa,IDA,,Angolan
kwanzas,1970,"MICS, 2003",MIS, 2008/07,,1997,,,IAP,1991-96,2003,BPM5,Actual,Special
,,GDDS,"IND, 2000",,1984-85,,1991,2000,AD,AD,Angola,Angola

ALB,Republic of Albania,Upper middle Income,Europe & Central Asia,IBRD,,Albanian Lek
,2001,"MICS, 2005",,1998,1993,VAB,,2005,BPM5,Actual,General,consolidated,GDDS,"LSMS
2005",Yes,1998,2003,2008,2000,AL,AL,Albania,Albania

ARE,United Arab Emirates,High Income: nonOECD,Middle East & North Africa,,U.A.E.
Dirham,2005,,,1995,,,VAB,,,BPM4,,General,consolidated,GDDS,,1998,,2008,2005,AE,AE
,United Arab Emirates,United Arab Emirates

ARG,Argentine Republic,Upper middle Income,Latin America & Caribbean,IBRD,,Argentine
peso,2003,,1991,,1991,VAB,1971-84,2005,BPM5,Actual,Special,Consolidated,SDDS,"IND,
2008",Yes,2002,2001,2008,2000,AR,AR,Argentina,Argentina
```

3 - The data looks like this when downloaded

CountryCode	Long Name	Income Group	Region	Lending
AWA	Aruba	High Income: nonOECD	Latin America & Caribbean	
ADO	Principality of Andorra	High Income: nonOECD	Europe & Central Asia	
AFG	Islamic State of Afghanistan	Low Income	South Asia	IDA
AGO	People's Republic of Angola	Lower middle Income	Sub-Saharan Africa	IDA
ALB	Republic of Albania	Upper middle Income	Europe & Central Asia	IBRD
ARE	United Arab Emirates	High Income: nonOECD	Middle East & North Africa	
ARG	Argentine Republic	Upper middle Income	Latin America & Caribbean	IBRD
ARM	Republic of Armenia	Lower middle Income	Europe & Central Asia	IBRD
ASM	American Samoa	Upper middle Income	Pacific Asia & Pacific	
ATG	Antigua and Barbuda	Upper middle Income	Latin America & Caribbean	IBRD
AUS	Commonwealth of Australia	High Income: OECD	Asia & Pacific	
AUT	Republic of Austria	High Income: OECD	Europe & Central Asia	
AZJ	Republic of Azerbaijan	Upper middle Income	Europe & Central Asia	IBRD
BDI	Republic of Burundi	Low Income	Sub-Saharan Africa	IDA
BEL	Kingdom of Belgium	High Income: OECD	Europe & Central Asia	
BRN	Republic of Brunei	Low Income	Sub-Saharan Africa	IDA
BFA	Burkina Faso	Low Income	Sub-Saharan Africa	IDA
BGD	People's Republic of Bangladesh	Low Income	South Asia	IDA

4 - After pre-processing

```
outcome <- tolower(outcome)
# column type is same as variable so changing it
chosen_state <- state
# Check that state and outcome are valid
if (length(unique(out_dt[,state])) > 0) &&
    length(unique(out_dt[,outcome])) > 0) {
  stop("invalid state")
}
if (length(unique(out_dt[,state])) > 0) &&
    length(unique(out_dt[,outcome])) > 0) {
  stop("invalid outcome")
}
# renaming columns to be less verbose and lowercase
setnames(out_dt,
         tolower(sapply(colnames(out_dt), gsub, pattern = "Noospital 30-day Death"))
)
# filter by state
out_dt <- out_dt[state == chosen_state]
# columns indices to keep
col_indices <- grep(paste0("hospital name|state|^",outcome), colnames(out_dt))
# filtering out unnecessary data
out_dt <- out_dt[, .SD[.SDcols = col_indices]]
# find out what class each column is
# sapply(out_dt[,col_indices], class)
out_dt[, outcome] <- out_dt[, as.numeric(get(outcome))]
# removing missing values for numerical datatype (outcome column)
out_dt <- out_dt[complete.cases(out_dt),]
# order column to top
```